and then washed with DMF and methylene chloride. Two unmasked cycles of photodeprotection and coupling with carboxy-activated NVOC-glycine were carried out. The surface was then illuminated through a chrome on glass 50 μm checkerboard pattern mask. Carboxy-activated *Nα*-tBOC-*O*-tButyl-L-tyrosine was then added. The entire surface was uniformly illuminated to photolyze the remaining NVOC groups. Finally, carboxy-activated NVOC-L-proline was added, the NVOC group was removed by illumination, and the t-BOC and t-butyl protecting groups were removed with TFA. After removal of the protecting groups, the surface consisted of a 50 μm checkerboard array of Tyr-Gly-Gly-Phe-Leu (YGGFL) and Pro-Gly-Gly-Phe-Leu (PGGFL). See also SEQ ID NO:1 and SEQ ID NO:2.

B. Antibody Recognition

In one preferred embodiment the substrate is used to determine which of a plurality of amino acid sequences is recognized by an antibody of interest.

1. Example

In one example, the array of pentapeptides in the example illustrated in Fig. 2 was probed with a mouse monoclonal antibody directed against β-endorphin. This antibody (called 3E7) is known to bind YGGFL and YGGFM (see also SEQ ID NO:1 and SEQ ID NO:21) with nanomolar affinity and is discussed in Meo et al., Proc. Natl. Acad. Sci. USA (1983) 80:4084, which is incorporated by reference herein for all purposes. This antibody requires the amino terminal tyrosine for high affinity binding. The array of peptides formed as described in Fig. 2 was incubated with a 2 μg/ml mouse monoclonal antibody (3E7) known to recognize YGGFL. See also SEQ ID NO:1. 3E7 does not bind PGGFL. See also SEQ ID NO:2. A second incubation with fluoresceinated goat anti-mouse antibody labeled the regions that bound 3E7. The surface was scanned with an epi-fluorescence microscope. The

results showed alternating bright and dark 50 μm squares indicating that YGGFL (SEQ ID NO:1) and PGGFL (SEQ ID NO:2) were synthesized in a geometric array determined by the mask. A high contrast (>12:1 intensity ratio) fluorescence checkerboard image shows that (a) YGGFL (SEQ ID NO:1) and PGGFL (SEQ ID NO:2) were synthesized in alternate 50 μm squares, (b) YGGFL (SEQ ID NO:1) attached to the surface is accessible for binding to antibody 3E7, and (c) antibody 3E7 does not bind to PGGFL (SEQ ID NO:2)

A three-dimensional representation of the fluorescence intensity data in a 2 square by 4 square rectangular portion of the checkerboard was produced. It shows that the border between synthesis sites is sharp. The height of each spike in this display is linearly proportional to the integrated fluorescence intensity in a 2.5 μm pixel. The transition between PGGFL and YGGFL occurs within two spikes (5 μm). There is little variation in the fluorescence intensity of different YGGFL squares. The mean intensity of sixteen YGGFL synthesis sites was 2.03×10^5 counts and the standard deviation was 9.6×10^3 counts.

III. Synthesis

A. Reactor System

Fig. 3 schematically illustrates a device used to synthesize diverse polymer sequences on a substrate. The device includes an automated peptide synthesizer 401. The automated peptide synthesizer is a device which flows selected reagents through a flow cell 402 under the direction of a computer 404. In a preferred embodiment the synthesizer is an ABI Peptide Synthesizer, model no. 431A. The computer may be selected from a wide variety of computers or discrete logic including for, example, an IBM PC-AT or similar computer linked with appropriate internal control systems in the peptide synthesizer. The PC is provided with signals from the

board computer indicative of, for example, the end of a coupling cycle.

Substrate 406 is mounted on the flow cell, forming a cavity between the substrate and the flow cell. Selected reagents flow through this cavity from the peptide synthesizer at selected times, forming an array of peptides on the face of the substrate in the cavity. Mounted above the substrate, and preferably in contact with the substrate is a mask 408. Mask 408 is transparent in selected regions to a selected wavelength of light and is opaque in other regions to the selected wavelength of light. The mask is illuminated with a light source 410 such as a UV light source. In one specific embodiment the light source 410 is a model no. 82420 made by Oriel. The mask is held and translated by an x-y-z translation stage 412 such as an x-y translation stage made by Newport Corp. The computer coordinates action of the peptide synthesizer, x-y translation stage, and light source. Of course, the invention may be used in some embodiments with translation of the substrate instead of the mask.

In operation, the substrate is mounted on the flow cell. The substrate, with its surface protected by a suitable photo removable protecting group, is exposed to light at selected locations by positioning the mask and directing light from a light source, through the mask, onto the substrate for a desired period of time (such as, for example, 1 sec to 60 min in the case of peptide synthesis). A selected peptide or other monomer/polymer is pumped through the reactor cavity by the peptide synthesizer for binding at the selected locations on the substrate. After a selected reaction time (such as about 1 sec to 300 min in the case of peptide reactions) the monomer is washed from the system, the mask is appropriately repositioned or replaced, and the cycle is repeated. In most embodiments of the

invention, reactions may be conducted at or near ambient temperature.

Figs 4a and 4b are flow charts of the software used in operation of the reactor system. At step 502 the peptide synthesis software is initialized. At step 504 the system calibrates positioners on the x-y translation stage and begins a main loop. At step 506 the system determines which, if any, of the function keys on the computer have been pressed. If F1 has been pressed, the system prompts the user for input of a desired synthesis process. If the user enters F2, the system allows a user to edit a file for a synthesis process at step 510. If the user enters F3 the system loads a process from a disk at step 512. If the user enters F4 the system saves an entered or edited process to disk at step 514. If the user selects F5 the current process is displayed at step 516 while selection of F6 starts the main portion of the program, i.e., the actual synthesis according to the selected process. If the user selects F7 the system displays the location of the synthesized peptides, while pressing F10 returns the user to the disk operating system.

Fig. 4b illustrates the synthesis step 518 in greater detail. The main loop of the program is started in which the system first moves the mask to a next position at step 526. During the main loop of the program, necessary chemicals flow through the reaction cell under the direction of the on-board computer in the peptide synthesizer. At step 528 the system then waits for an exposure command and, upon receipt of the exposure command exposes the substrate for a desired time at step 530. When an acknowledgement of complete exposure is received at step 532 the system determines if the process is complete at step 534 and, if so, waits for additional keyboard input at step 536 and, thereafter, exits the perform synthesis process.

A computer program used for operation of the system described above is written in Turbo C (Borland Int'l) and has been implemented in an IBM compatible system. The motor control software is adapted from software produced by Newport Corporation. It will be recognized that a large variety of programming languages could be utilized without departing from the scope of the invention herein. Certain calls are made to a graphics program in "Programmer Guide to PC and PS2 Video Systems" (Wilton, Microsoft Press, 1987), which is incorporated herein by reference for all purposes.

Alignment of the mask is achieved by one of two methods in preferred embodiments. In a first embodiment the system relies upon relative alignment of the various components, which is normally acceptable since x-y-z translation stages are capable of sufficient accuracy for the purposes herein. In alternative embodiments, alignment marks on the substrate are coupled to a CCD device for appropriate alignment.

According to some embodiments, pure reagents are not added at each step, or complete photolysis of the protecting groups is not provided at each step. According to these embodiments, multiple products will be formed in each synthesis site. For example, if the monomers A and B are mixed during a synthesis step, A and B will bind to deprotected regions, roughly in proportion to their concentration in solution. Hence, a mixture of compounds will be formed in a synthesis region. A substrate formed with mixtures of compounds in various synthesis regions may be used to perform, for example, an initial screening of a large number of compounds, after which a smaller number of compounds in regions which exhibit high binding affinity are further screened. Similar results may be obtained by only partially photolyzing a region, adding a first monomer, re-photolyzing the same region, and exposing the region to a second monomer.

B. Binary Synthesis Strategy

In a light-directed chemical synthesis, the products formed depend on the pattern and order of masks, and on the order of reactants. To make a set of products there will in general be a finite number of possible masking strategies. In preferred embodiments of the invention herein a binary synthesis strategy is utilized. The binary synthesis strategy is illustrated herein primarily with regard to a masking strategy, although it will be applicable to other polymer synthesis strategies such as the pin strategy, and the like.

In a binary synthesis strategy, the substrate is irradiated with a first mask, exposed to a first building block, irradiated with a second mask, exposed to a second building block, etc. Each combination of masked irradiation and exposure to a building block is referred to herein as a "cycle."

In a preferred binary masking strategy, the masks for each cycle allow illumination of half of a region of interest on the substrate and no illumination of the remaining half of the region of interest. By "half" it is intended herein not to mean exactly one-half the region of interest, but instead a large fraction of the region of interest such as from about 30 to 70 percent of the region of interest. It will be understood that the entire masking strategy need not take a binary form; instead non-binary cycles may be introduced as desired between binary cycles.

In preferred embodiments of the binary masking strategy, a given cycle illuminates only about half of the region which was illuminated in a previous cycle, while not illuminating the remaining half of the illuminated portion from the previous cycle. Conversely, in such preferred embodiments, a given cycle illuminates half of the region which was not illuminated in the

previous cycle and does not illuminate half the region which was not illuminated in a previous cycle.

In the synthesis strategy disclosed herein, the longest length (ℓ) of the synthesized polymers is $\ell = n/a$; where n is the number of cycles and a is the number of chemical building blocks (note that a given building block may be repeated).

The synthesis strategy is most readily illustrated and handled in matrix notation. At each synthesis site, the determination of whether to add a given monomer is a binary process. Therefore, each product element P_j is given by the dot product of two vectors, a chemical reactant vector, e.g., C = [A,B,C,D], and a binary vector σ_j. Inspection of the products in the example below for a four-step synthesis, shows that in one four-step synthesis σ_1 = [1,0,1,0], σ_2 = [1,0,0,1], σ_3 = [0,1,1,0], and σ_4 = [0,1,0,1], where a 1 indicates illumination and a 0 indicates no illumination. Therefore, it becomes possible to build a "switch matrix" S from the column vectors σ_j (j = 1,k where k is the number of products).

$$S = \begin{matrix} \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{matrix}$$

The outcome P of a synthesis is simply P = CS, the product of the chemical reactant matrix and the switch matrix.

The switch matrix for an n-cycle synthesis yielding k products has n rows and k columns. An important attribute of S is that each row specifies a mask. A two-dimensional mask m_j for the jth chemical step of a synthesis is obtained directly from the jth row of S by placing the elements $s_{j1},\ldots s_{jk}$ into, for example, a

square format. The particular arrangement below provides a square format, although linear or other arrangements may be utilized.

$$S = \begin{matrix} s_{11} & s_{12} & s_{13} & s_{14} \\ s_{21} & s_{22} & s_{23} & s_{24} \\ s_{31} & s_{32} & s_{33} & s_{34} \\ s_{41} & s_{42} & s_{43} & s_{44} \end{matrix} \qquad m_j = \begin{matrix} s_{j1} & s_{j2} \\ s_{j3} & s_{j4} \end{matrix}$$

Of course, compounds formed in a light-activated synthesis can be positioned in any defined geometric array. A square or rectangular matrix is convenient but not required. The rows of the switch matrix may be transformed into any convenient array as long as equivalent transformations are used for each row.

For example, the masks in the four-step synthesis below are then denoted by:

$$m_1 = \begin{matrix} 1 & 1 \\ 0 & 0 \end{matrix} \qquad m_2 = \begin{matrix} 0 & 0 \\ 1 & 1 \end{matrix} \qquad m_3 = \begin{matrix} 1 & 0 \\ 1 & 0 \end{matrix} \qquad m_4 = \begin{matrix} 0 & 1 \\ 0 & 1 \end{matrix}$$

where 1 denotes illumination (activation) and 0 denotes no illumination.

The matrix representation is used to generate a desired set of products and product maps in preferred embodiments. Each compound is defined by the product of the chemical vector and a particular switch vector. Therefore, for each synthesis address, one simply saves the switch vector, assembles all of them into a switch matrix, and extracts each of the rows to form the masks.

In some cases, particular product distributions or a maximal number of products are desired. For example, for C = [A,B,C,D], any switch vector (σ_j) consists of four bits. Sixteen four-bit vectors exist. Hence, a maximum of 16 different products can be made by sequential addition of the reagents [A,B,C,D]. These 16 column vectors can be assembled in 16! different ways to

form a switch matrix. The order of the column vectors defines the masking patterns, and therefore, the spatial ordering of products but not their makeup. One ordering of these columns gives the following switch matrix (in which "null" (∅) additions are included in brackets for the sake of completeness, although such null additions are elsewhere ignored herein):

	$\sigma 1$															σ_{16}	
	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	A
	[0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1]	∅
S =	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	B
	[0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1]	∅
	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	C
	[0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1]	∅
	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	D
	[0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1]	∅

The columns of S according to this aspect of the invention are the binary representations of the numbers 15 to 0. The sixteen products of this binary synthesis are ABCD, ABC, ABD, AB, ACD, AC, AD, A, BCD, BC, BD, B, CD, C, D, and ∅ (null). Also note that each of the switch vectors from the four-step synthesis masks above (and hence the synthesis products) are present in the four bit binary switch matrix. (See columns 6, 7, 10, and 11)

This synthesis procedure provides an easy way for mapping the completed products. The products in the various locations on the substrate are simply defined by the columns of the switch matrix (the first column indicating, for example, that the product ABCD will be present in the upper left-hand location of the substrate). Furthermore, if only selected desired products are to be made, the mask sequence can be derived by extracting the columns with the desired sequences. For example, to form the product set ABCD, ABD, ACD, AD,

BCD, BD, CD, and D, the masks are formed by use of a switch matrix with only the 1st, 3rd, 5th, 7th, 9th, 11th, 13th, and 15th columns arranged into the switch matrix:

$$S = \begin{matrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{matrix}$$

To form all of the polymers of length 4, the reactant matrix [ABCDABCDABCDABCD] is used. The switch matrix will be formed from a matrix of the binary numbers from 0 to 2^{16} arranged in columns. The columns having four monomers are then selected and arranged into a switch matrix. Therefore, it is seen that the binary switch matrix in general will provide a representation of all the products which can be made from an n-step synthesis, from which the desired products are then extracted.

The rows of the binary switch matrix will, in preferred embodiments, have the property that each masking step illuminates half of the synthesis area. Each masking step also factors the preceding masking step; that is, half of the region that was illuminated in the preceding step is again illuminated, whereas the other half is not. Half of the region that was not illuminated in the preceding step is also illuminated, whereas the other half is not. Thus, masking is recursive. The masks are constructed, as described previously, by extracting the elements of each row and placing them in a square array. For example, the four masks in S for a four-step synthesis are:

$$m_1 = \begin{matrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} \quad m_2 = \begin{matrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{matrix} \quad m_3 = \begin{matrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{matrix} \quad m_4 = \begin{matrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{matrix}$$

The recursive factoring of masks allows the products of a light-directed synthesis to be represented by a polynomial. (Some light activated syntheses can only be denoted by irreducible, i.e., prime polynomials.) For example, the polynomial corresponding to the top synthesis of Fig. 8a (discussed below) is

$$P = (A + B)(C + D)$$

A reaction polynomial may be expanded as though it were an algebraic expression, provided that the order of joining of reactants X_1 and X_2 is preserved ($X_1X_2 \neq X_2X_1$), i.e., the products are not commutative. The product then is AC + AD + BC + BD. The polynomial explicitly specifies the reactants and implicitly specifies the mask for each step. Each pair of parentheses demarcates a round of synthesis. The chemical reactants of a round (e.g., A and B) react at nonoverlapping sites and hence cannot combine with one another. The synthesis area is divided equally amongst the elements of a round (e.g., A is directed to one-half of the area and B to the other half). Hence, the masks for a round (e.g., the masks m_A and m_B) are orthogonal and form an orthonormal set. The polynomial notation also signifies that each element in a round is to be joined to each element of the next round (e.g., A with C, A with D, B with C, and B with D). This is accomplished by having m_C overlap m_A and m_B equally, and likewise for m_D. Because C and D are elements of a round, m_C and m_D are orthogonal to each other and form an orthonormal set.

The polynomial representation of the binary synthesis described above, in which 16 products are made from 4 reactants, is

$$P = (A + \emptyset)\ (B + \emptyset)\ (C + \emptyset)\ (D + \emptyset)$$

which gives ABCD, ABC, ABD, AB, ACD, AC, AD, A, BCD, BC, BD, B, CD, C, D, and Ø when expanded (with the rule that ØX = X and XØ = X, and remembering that joining is ordered). In a binary synthesis, each round contains one reactant and one null (denoted by Ø). Half of the synthesis area receives the reactant and the other half receives nothing. Each mask overlaps every other mask equally.

Binary rounds and non-binary rounds can be interspersed as desired, as in

$$P = (A + \emptyset)\ (B)\ (C + D + \emptyset)\ (E + F + G)$$

The 18 compounds formed are ABCE, ABCF, ABCG, ABDE, ABDF, ABDG, ABE, ABF, ABG, BCE, BCF, BCG, BDE, BDF, BDG, BE, BF, and BG. The switch matrix S for this 7-step synthesis is

$$S = \begin{matrix} 1&1&1&1&1&1&1&1&1&0&0&0&0&0&0&0&0&0 \\ 1&1&1&1&1&1&1&1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&0&0&0&0&0&0&1&1&1&0&0&0&0&0&0 \\ 0&0&0&1&1&1&0&0&0&0&0&0&1&1&1&0&0&0 \\ 1&0&0&1&0&0&1&0&0&1&0&0&1&0&0&1&0&0 \\ 0&1&0&0&1&0&0&1&0&0&1&0&0&1&0&0&1&0 \\ 0&0&1&0&0&1&0&0&1&0&0&1&0&0&1&0&0&1 \end{matrix}$$

The round denoted by (B) places B in all products because the reaction area was uniformly activated (the mask for B consisted entirely of 1's).

The number of compounds k formed in a synthesis consisting of r rounds, in which the ith round has b_i chemical reactants and z_i nulls, is

$$k = \Sigma\ (b_i + z_i)$$

and the number of chemical steps n is

$$n = \Sigma b_i$$

The number of compounds synthesized when b = a (the number of chemical building blocks) and z = 0 in all rounds is $a^{n/a}$, compared with 2^n for a binary synthesis. For n = 20 and a = 5, 625 compounds (all tetramers) would be formed, compared with 1.049×10^6 compounds in a binary synthesis with the same number of chemical steps.

It should also be noted that rounds in a polynomial can be nested, as in

$$(A + (B+\emptyset)(C+\emptyset))\ (D+\emptyset)$$

The products are AD, BCD, BD, CD, D, A, BC, B, C, and $\emptyset$.

Binary syntheses are attractive for two reasons. First, they generate the maximal number of products (2^n) for a given number of chemical steps (n). For four reactants, 16 compounds are formed in the binary synthesis, whereas only 4 are made when each round has two reactants. A 10-step binary synthesis yields 1,024 compounds, and a 20-step synthesis yields 1,048,576. Second, products formed in a binary synthesis are a complete nested set with lengths ranging from 0 to n. All compounds that can be formed by deleting one or more units from the longest product (the n-mer) are present. Contained within the binary set are the smaller sets that would be formed from the same reactants using any other set of masks (e.g., AC, AD, BC, and BD formed in the synthesis shown in Fig. 5 are present in the set of 16 formed by the binary synthesis). In some cases, however, the experimentally achievable spatial resolution may not suffice to accommodate all the compounds formed. Therefore, practical limitations may require one to select a particular subset of the possible switch vectors for a given synthesis.

1. Example

Fig. 5 illustrates a synthesis with a binary masking strategy. The binary masking strategy provides the greatest number of sequences for a given number of cycles. According to this embodiment, a mask m_1 allows illumination of half of the substrate. The substrate is then exposed to the building block A, which binds at the illuminated regions.

Thereafter, the mask m_2 allows illumination of half of the previously illuminated region, while it does not illuminate half of the previously illuminated region. The building block B is then added, which binds at the illuminated regions from m_2.

The process continues with masks m_3, m_4, and m_5, resulting in the product array shown in the bottom portion of the figure. The process generates 32 (2 raised to the power of the number of monomers) sequences with 5 (the number of monomers) cycles.

2. Example

Fig. 6 illustrates another preferred binary masking strategy which is referred to herein as the gray code masking strategy. According to this embodiment, the masks m_1 to m_5 are selected such that a side of any given synthesis region is defined by the edge of only one mask. The site at which the sequence BCDE is formed, for example, has its right edge defined by m_5 and its left side formed by mask m_4 (and no other mask is aligned on the sides of this site). Accordingly, problems created by misalignment, diffusion of light under the mask and the like will be minimized.

3. Example

Fig. 7 illustrates another binary masking strategy. According to this scheme, referred to herein as a modified gray code masking strategy, the number of masks needed is minimized. For example, the mask m_2 could

be the same mask as m1 and simply translated laterally. Similarly, the mask m_4 could be the same as mask m_3 and simply translated laterally.

4. Example

A four-step synthesis is shown in Fig. 8a. The reactants are the ordered set {A,B,C,D}. In the first cycle, illumination through m_1 activates the upper half of the synthesis area. Building block A is then added to give the distribution 602. Illumination through mask m_2 (which activates the lower half), followed by addition of B yields the next intermediate distribution 604. C is added after illumination through m_3 (which activates the left half) giving the distribution 604, and D after illumination through m_4 (which activates the right half), to yield the final product pattern 608 {AC,AD,BC,BD}.

5. Example

The above masking strategy for the synthesis may be extended for all 400 dipeptides from the 20 naturally occurring amino acids as shown in Fig. 8b. The synthesis consists of two rounds, with 20 photolysis and chemical coupling cycles per round. In the first cycle of round 1, mask 1 activates 1/20th of the substrate for coupling with the first of 20 amino acids. Nineteen subsequent illumination/coupling cycles in round 1 yield a substrate consisting of 20 rectangular stripes each bearing a distinct member of the 20 amino acids. The masks of round 2 are perpendicular to round 1 masks and therefore a single illumination/coupling cycle in round 2 yields 20 dipeptides. The 20 illumination/coupling cycles of round 2 complete the synthesis of the 400 dipeptides.

6. Example

The power of the binary masking strategy can be appreciated by the outcome of a 10-step synthesis that

produced 1,024 peptides. The polynomial expression for this 10-step binary synthesis was:

(f+Ø) (Y+Ø) (G+Ø) (A+Ø) (G+Ø) (T+Ø) (F+Ø) (L+Ø) (S+Ø) (F+Ø)

Each peptide occupied a 400x400 μm square. A 32x32 peptide array (1,024 peptides, including the null peptide and 10 peptides of $\ell = 1$, and a limited number of duplicates) was clearly evident in a fluorescence scan following side group deprotection and treatment with the antibody 3E7 and fluoresceinated antibody. Each synthesis site was a 400x400 μm square.

The scan showed a range of fluorescence intensities, from a background value of 3,300 counts to 22,400 counts in the brightest square ($x = 20$, $y = 9$). Only 15 compounds exhibited an intensity greater than 12,300 counts. The median value of the array was 4,800 counts.

The identity of each peptide in the array could be determined from its x and y coordinates (each range from 0 to 31) and the map of Fig. 9. The chemical units at positions 2, 5, 6, 9, and 10 are specified by the y coordinate and those at positions 1, 3, 4, 7, 8 by the x coordinate. All but one of the peptides was shorter than 10 residues. For example, the peptide at $x = 12$ and $y = 3$ is YGAGF (SEQ ID NO:3; positions 1, 6, 8, 9, and 10 are nulls). YGAFLS (SEQ ID NO:4), the brightest element of the array, is at $x = 20$ and $y = 9$.

It is often desirable to deduce a binding affinity of a given peptide from the measured fluorescence intensity. Conceptually, the simplest case is one in which a single peptide binds to a univalent antibody molecule. The fluorescence scan is carried out after the slide is washed with buffer for a defined time. The order of fluorescence intensities is then a measure primarily of the relative dissociation rates of the antibody-peptide complexes. If the on-rate constants are

the same (e.g., if they are diffusion-controlled), the order of fluorescence intensities will typically correspond to the order of binding affinities. However, the situation is sometimes more complex because a bivalent primary antibody and a bivalent secondary antibody are used. The density of peptides in a synthesis area corresponded to a mean separation of ~7 nm, which would allow multivalent antibody-peptide interactions. Hence, fluorescence intensities obtained according to the method herein will often be a qualitative indicator of binding affinity.

Another important consideration is the fidelity of synthesis. Deletions are produced by incomplete photodeprotection or incomplete coupling. The coupling yield per cycle in these experiments is typically between 85% and 95%. Implementing the switch matrix by masking is imperfect because of light diffraction, internal reflection, and scattering. Consequently, stowaways (chemical units that should not be on board) arise by unintended illumination of regions that should be dark. A binary synthesis array contains many of the controls needed to assess the fidelity of a synthesis. For example, the fluorescence signal from a synthesis area nominally containing a tetrapeptide ABCD could come from a tripeptide deletion impurity such as ACD. Such an artifact would be ruled out by the finding that the fluorescence intensity of the ACD site is less than that of the ABCD site.

The fifteen most highly fluorescent peptides in the array obtained with the synthesis of 1,024 peptides described above, were YGAFLS (SEQ ID NO:4), YGAFS (SEQ ID NO:5), YGAFL (SEQ ID NO:6), YGGFLS (SEQ ID NO:7), YGAF (SEQ ID NO:8), YGALS (SEQ ID NO:9), YGGFS (SEQ ID NO:10), YGAL (SEQ ID NO:11), YGAFLF (SEQ ID NO:12), YGAF (SEQ ID NO:8), YGAFF (SEQ ID NO:13), YGGLS (SEQ ID NO:14), YGGFL (SEQ ID NO:1 and SEQ ID NO:15), YGAFSF (SEQ ID NO:16), and YGAFLSF (SEQ ID NO:17). A striking feature is that

all fifteen begin with YG, which agrees with previous work showing that an amino-terminal tyrosine is a key determinant of binding to 3E7. Residue 3 of this set is either A or G, and residue 4 is either F or L. The exclusion of S and T from these positions is clear cut. The finding that the preferred sequence is YG (A/G) (F/L) fits nicely with the outcome of a study in which a very large library of peptides on phage generated by recombinant DNA methods was screened for binding to antibody 3E7 (see Cwirla et al., Proc. Natl. Acad. Sci. USA, (1990) 87:6378, incorporated herein by reference). Additional binary syntheses based on leads from peptides on phage experiments show that YGAFMQ (SEQ ID NO:18), YGAFM (SEQ ID NO:19), and YGAFQ (SEQ ID NO:20) give stronger fluorescence signals than does YGGFM (SEQ ID NO:21), the immunogen used to obtain antibody 3E7.

Variations on the above masking strategy will be valuable in certain circumstances. For example, if a "kernel" sequence of interest consists of PQR separated from XYZ, the aim is to synthesize peptides in which these units are separated by a variable number of different residues. The kernel can be placed in each peptide by using a mask that has 1's everywhere. The polynomial representation of a suitable synthesis is:

(P) (Q) (R) (A+∅) (B+∅) (C+∅) (D+∅) (X) (Y) (Z)

Sixteen peptides will be formed, ranging in length from the 6-mer PQRXYZ to the 10-mer PQRABCDXYZ.

Several other masking strategies will also find value in selected circumstances. By using a particular mask more than once, two or more reactants will appear in the same set of products. For example, suppose that the mask for an 8-step synthesis is

A 11110000
B 00001111

C 11001100
D 00110011
E 10101010
F 01010101
G 11110000
H 00001111

The products are ACEG, ACFG, ADEG, ADFG, BCEH, BCFH, BDEH, and BDFH. A and G always appear in the same product, although not necessarily next to each other, because their additions were directed by the same mask, and likewise for B and H.

C. Linker Selection

According to preferred embodiments the linker molecules used as an intermediary between the synthesized polymers and the substrate are selected for optimum length and/or type for improved binding interaction with a receptor. According to this aspect of the invention diverse linkers of varying length and/or type are synthesized for subsequent attachment of a ligand. Through variations in the length and type of linker, it becomes possible to optimize the binding interaction between an immobilized ligand and its receptor.

The degree of binding between a ligand (peptide, inhibitor, hapten, drug, etc.) and its receptor (enzyme, antibody, etc.) when one of the partners is immobilized on to a substrate will in some embodiments depend on the accessibility of the receptor in solution to the immobilized ligand. The accessibility in turn will depend on the length and/or type of linker molecule employed to immobilize one of the partners. Preferred embodiments of the invention therefore employ the VLSIPS technology described herein to generate an array of, preferably, inactive or inert linkers of varying length and/or type, using photochemical protecting groups to

selectively expose different regions of the substrate and to build upon chemically-active groups.

In the simplest embodiment of this concept, the same unit is attached to the substrate in varying multiples or lengths in known locations on the substrate via VLSIPS techniques to generate an array of polymers of varying length. A single ligand (peptide, drug, hapten, etc.) is attached to each of them, and an assay is performed with the binding site to evaluate the degree of binding with a receptor that is known to bind to the ligand. In cases where the linker length impacts the ability of the receptor to bind to the ligand, varying levels of binding will be observed. In general, the linker which provides the highest binding will then be used to assay other ligands synthesized in accordance with the techniques herein.

According to other embodiments the binding between a single ligand/receptor pair is evaluated for linkers of diverse monomer sequence. According to these embodiments, the linkers are synthesized in an array in accordance with the techniques herein and have different monomer sequences (and, optionally, different lengths). Thereafter, all of the linker molecules are provided with a ligand known to have at least some binding affinity for a given receptor. The given receptor is then exposed to the ligand and binding affinity is deduced. Linker molecules which provide adequate binding between the ligand and receptor are then utilized in screening studies.

D. Protecting Groups

As discussed above, selectively removable protecting groups allow creation of well defined areas of substrate surface having differing reactivities. Preferably, the protecting groups are selectively removed from the surface by applying a specific activator, such as electromagnetic radiation of a specific wavelength and

intensity. More preferably, the specific activator exposes selected areas of the surface to remove the protecting groups in the exposed areas.

Protecting groups of the present invention are used in conjunction with solid phase oligomer syntheses, such as peptide syntheses using natural or unnatural amino acids, nucleotide syntheses using deoxyribonucleic and ribonucleic acids, oligosaccharide syntheses, and the like. In addition to protecting the substrate surface from unwanted reaction, the protecting groups block a reactive end of the monomer to prevent self-polymerization. For instance, attachment of a protecting group to the amino terminus of an activated amino acid, such as an N-hydroxysuccinimide-activated ester of the amino acid, prevents the amino terminus of one monomer from reacting with the activated ester portion of another during peptide synthesis. Alternatively, the protecting group may be attached to the carboxyl group of an amino acid to prevent reaction at this site. Most protecting groups can be attached to either the amino or the carboxyl group of an amino acid, and the nature of the chemical synthesis will dictate which reactive group will require a protecting group. Analogously, attachment of a protecting group to the 5'-hydroxyl group of a nucleoside during synthesis using for example, phosphate-triester coupling chemistry, prevents the 5'-hydroxyl of one nucleoside from reacting with the 3'-activated phosphate-triester of another.

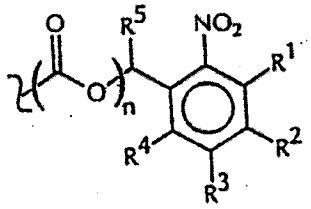
Regardless of the specific use, protecting groups are employed to protect a moiety on a molecule from reacting with another reagent. Protecting groups of the present invention have the following characteristics: they prevent selected reagents from modifying the group to which they are attached; they are stable (that is, they remain attached to the molecule) to the synthesis reaction conditions; they are removable under conditions that do not adversely affect the remaining structure; and

once removed, they do not react appreciably with the surface or surface-bound oligomer. The selection of a suitable protecting group will depend, of course, on the chemical nature of the monomer unit and oligomer, as well as the specific reagents they are to protect against.

In a preferred embodiment, the protecting groups are photoactivatable. The properties and uses of photoreactive protecting compounds have been reviewed. See, McCray et al., Ann. Rev. of Biophys. and Biophys. Chem. (1989) 18:239-270, which is incorporated herein by reference. Preferably, the photosensitive protecting groups will be removable by radiation in the ultraviolet (UV) or visible portion of the electromagnetic spectrum. More preferably, the protecting groups will be removable by radiation in the near UV or visible portion of the spectrum. In some embodiments, however, activation may be performed by other methods such as localized heating, electron beam lithography, laser pumping, oxidation or reduction with microelectrodes, and the like. Sulfonyl compounds are suitable reactive groups for electron beam lithography. Oxidative or reductive removal is accomplished by exposure of the protecting group to an electric current source, preferably using microelectrodes directed to the predefined regions of the surface which are desired for activation. Other methods may be used in light of this disclosure.

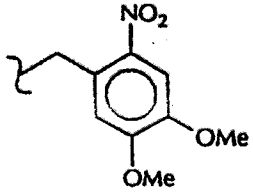
Many, although not all, of the photoremovable protecting groups will be aromatic compounds that absorb near-UV and visible radiation. Suitable photoremovable protecting groups are described in, for example, McCray et al., Patchornik, J. Amer. Chem. Soc. (1970) 92:6333, and Amit et al., J. Org. Chem. (1974) 39:192, which are incorporated herein by reference.

A preferred class of photoremovable protecting groups has the general formula:

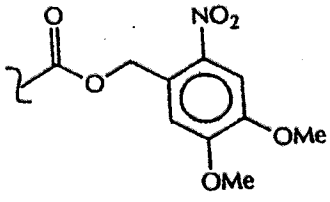


where R^1, R^2, R^3, and R^4 independently are a hydrogen atom, a lower alkyl, aryl, benzyl, halogen, hydroxyl, alkoxyl, thiol, thioether, amino, nitro, carboxyl, formate, formamido or phosphido group, or adjacent substituents (i.e., R^1-R^2, R^2-R^3, R^3-R^4) are substituted oxygen groups that together form a cyclic acetal or ketal; R^5 is a hydrogen atom, a alkoxyl, alkyl, halo, aryl, or alkenyl group, and n = 0 or 1.

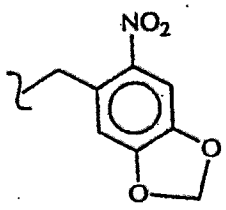
A preferred protecting group, 6-nitroveratryl (NV), which is used for protecting the carboxyl terminus of an amino acid or the hydroxyl group of a nucleotide, for example, is formed when R^2 and R^3 are each a methoxy group, R^1, R^4 and R^5 are each a hydrogen atom, and n = 0:



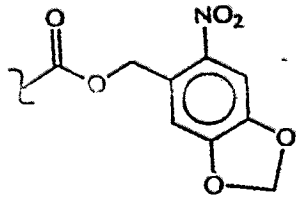
A preferred protecting group, 6-nitroveratryloxycarbonyl (NVOC), which is used to protect the amino terminus of an amino acid, for example, is formed when R^2 and R^3 are each a methoxy group, R^1, R^4 and R^5 are each a hydrogen atom, and n = 1:



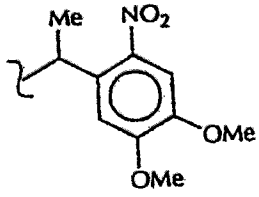
Another preferred protecting group, 6-nitropiperonyl (NP), which is used for protecting the carboxyl terminus of an amino acid or the hydroxyl group of a nucleotide, for example, is formed when R^2 and R^3 together form a methylene acetal, R^1, R^4 and R^5 are each a hydrogen atom, and n = 0:



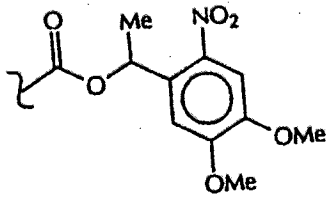
Another preferred protecting group, 6-nitropiperonyloxycarbonyl (NPOC), which is used to protect the amino terminus of an amino acid, for example, is formed when R^2 and R^3 together form a methylene acetal, R^1, R^4 and R^5 are each a hydrogen atom, and n = 1:



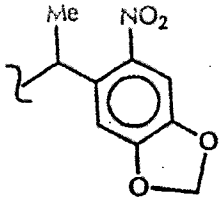
A most preferred protecting group, methyl-6-nitroveratryl (MeNV), which is used for protecting the carboxyl terminus of an amino acid or the hydroxyl group of a nucleotide, for example, is formed when R^2 and R^3 are each a methoxy group, R^1 and R^4 are each a hydrogen atom, R^5 is a methyl group, and n = 0:



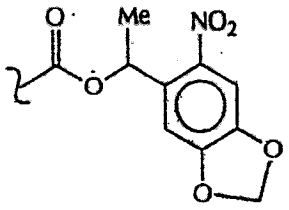
Another most preferred protecting group, methyl-6-nitroveratryloxycarbonyl (MeNVOC), which is used to protect the amino terminus of an amino acid, for example, is formed when R^2 and R^3 are each a methoxy group, R^1 and R^4 are each a hydrogen atom, R^5 is a methyl group, and n = 1:



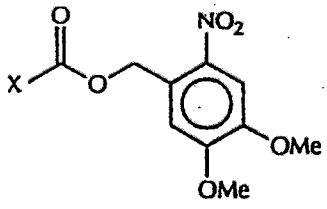
Another most preferred protecting group, methyl-6-nitropiperonyl (MeNP), which is used for protecting the carboxyl terminus of an amino acid or the hydroxyl group of a nucleotide, for example, is formed when R^2 and R^3 together form a methylene acetal, R^1 and R^4 are each a hydrogen atom, R^5 is a methyl group, and n = 0:



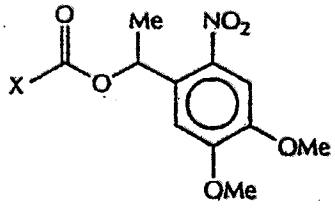
Another most preferred protecting group, methyl-6-nitropiperonyloxycarbonyl (MeNPOC), which is used to protect the amino terminus of an amino acid or to protect te 5' hydroxyl of nucleosides, for example, is formed when R^2 and R^3 together form a methylene acetal, R^1 and R^4 are each a hydrogen atom, R^5 is a methyl group, and n = 1:



A protected amino acid having a photoactivatable oxycarbonyl protecting group, such NVOC or NPOC or their corresponding methyl derivatives, MeNVOC or MeNPOC, respectively, on the amino terminus is formed by acylating the amine of the amino acid with an activated oxycarbonyl ester of the protecting group. Examples of activated oxycarbonyl esters of NVOC and MeNVOC have the general formula:



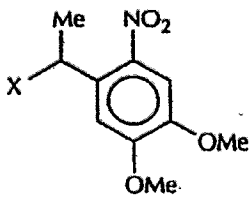
NVOC-X



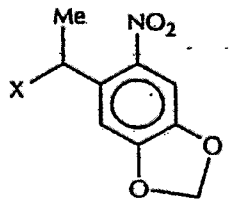
MeNVOC-X

where X is halogen, mixed anhydride, phenoxy, p-nitrophenoxy, N-hydroxysuccinimide, and the like.

A protected amino acid or nucleotide having a photoactivatable protecting group, such as NV or NP or their corresponding methyl derivatives, MeNV or MeNP, respectively, on the carboxy terminus of the amino acid or 5'-hydroxy terminus of the nucleotide, is formed by acylating the carboxy terminus or 5'-OH with an activated benzyl derivative of the protecting group. Examples of activated benzyl derivatives of MeNV and MeNP have the general formula:



MeNV-X

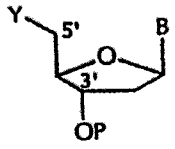


MeNP-X

where X is halogen, hydroxyl, tosyl, mesyl, trifluoromethyl, diazo, azido, and the like.

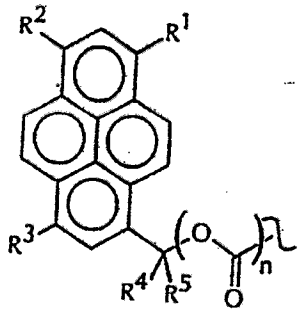
Another method for generating protected monomers is to react the benzylic alcohol derivative of the protecting group with an activated ester of the monomer. For example, to protect the carboxyl terminus of an amino acid, an activated ester of the amino acid is reacted with the alcohol derivative of the protecting group, such as 6-nitroveratrol (NVOH). Examples of activated esters suitable for such uses include halo-formate, mixed anhydride, imidazoyl formate, acyl halide, and also include formation of the activated ester in situ the use of common reagents such as DCC and the like. See Atherton et al. for other examples of activated esters.

A further method for generating protected monomers is to react the benzylic alcohol derivative of the protecting group with an activated carbon of the monomer. For example, to protect the 5'-hydroxyl group of a nucleic acid, a derivative having a 5'-activated carbon is reacted with the alcohol derivative of the protecting group, such as methyl-6-nitropiperonol (MePyROH). Examples of nucleotides having activating groups attached to the 5'-hydroxyl group have the general formula:



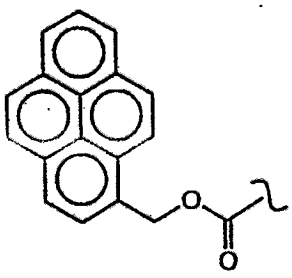
where Y is a halogen atom, a tosyl, mesyl, trifluoromethyl, azido, or diazo group, and the like.

Another class of preferred photochemical protecting groups has the formula:

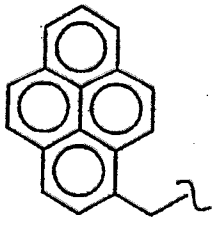


where R^1, R^2, and R^3 independently are a hydrogen atom, a lower alkyl, aryl, benzyl, halogen, hydroxyl, alkoxyl, thiol, thioether, amino, nitro, carboxyl, formate, formamido, sulfanates, sulfido or phosphido group, R^4 and R^5 independently are a hydrogen atom, an alkoxy, alkyl, halo, aryl, or alkenyl group, and n = 0 or 1.

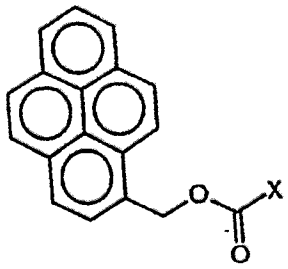
A preferred protecting group, 1-pyrenylmethyloxycarbonyl (PyROC), which is used to protect the amino terminus of an amino acid, for example, is formed when R^1 through R^5 are each a hydrogen atom and n = 1:



Another preferred protecting group, 1-pyrenylmethyl (PyR), which is used for protecting the carboxy terminus of an amino acid or the hydroxyl group of a nucleotide, for example, is formed when R^1 through R^5 are each a hydrogen atom and n = 0:

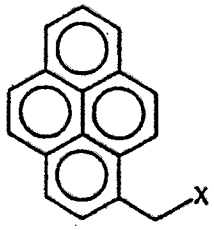


An amino acid having a pyrenylmethyloxycarbonyl protecting group on its amino terminus is formed by acylation of the free amine of amino acid with an activated oxycarbonyl ester of the pyrenyl protecting group. Examples of activated oxycarbonyl esters of PyROC have the general formula:



where X is halogen, or mixed anhydride, p-nitrophenoxy, or N-hydroxysuccinimide group, and the like.

A protected amino acid or nucleotide having a photoactivatable protecting group, such as PyR, on the carboxy terminus of the amino acid or 5'-hydroxy terminus of the nucleic acid, respectively, is formed by acylating the carboxy terminus or 5'-OH with an activated pyrenylmethyl derivative of the protecting group. Examples of activated pyrenylmethyl derivatives of PyROC have the general formula:



where X is a halogen atom, a hydroxyl, diazo, or azido group, and the like.

Another method of generating protected monomers is to react the pyrenylmethyl alcohol moiety of the protecting group with an activated ester of the monomer.

For example, an activated ester of an amino acid can be reacted with the alcohol derivative of the protecting group, such as pyrenylmethyl alcohol (PyROH), to form the protected derivative of the carboxy terminus of the amino acid. Examples of activated esters include halo-formate, mixed anhydride, imidazoyl formate, acyl halide, and also include formation of the activated ester in situ and the use of common reagents such as DCC and the like.

Clearly, many photosensitive protecting groups are suitable for use in the present invention.

In preferred embodiments, the substrate is irradiated to remove the photoremovable protecting groups and create regions having free reactive moieties and side products resulting from the protecting group. The removal rate of the protecting groups depends on the wavelength and intensity of the incident radiation, as well as the physical and chemical properties of the protecting group itself. Preferred protecting groups are removed at a faster rate and with a lower intensity of radiation. For example, at a given set of conditions, MeNVOC and MeNPOC are photolytically removed from the N-terminus of a peptide chain faster than their unsubstituted parent compounds, NVOC and NPOC, respectively.

Removal of the protecting group is accomplished by irradiation to separate the reactive group and the degradation products derived from the protecting group. Not wishing to be bound by theory, it is believed that irradiation of an NVOC- and MeNVOC-protected oligomers occurs by the following reaction schemes:

NVOC-AA -> 3,4-dimethoxy-6-nitrosobenzaldehyde + CO_2 + AA
MeNVOC-AA-> 3,4-dimethoxy-6-nitrosoacetophenone + CO_2 + AA

where AA represents the N-terminus of the amino acid oligomer.

Along with the unprotected amino acid, other products are liberated into solution: carbon dioxide and a 2,3-dimethoxy-6-nitrosophenylcarbonyl compound, which can react with nucleophilic portions of the oligomer to form unwanted secondary reactions. In the case of an NVOC-protected amino acid, the degradation product is a nitrosobenzaldehyde, while the degradation product for the other is a nitrosophenyl ketone. For instance, it is believed that the product aldehyde from NVOC degradation reacts with free amines to form a Schiff base (imine) that affects the remaining polymer synthesis. Preferred photoremovable protecting groups react slowly or reversibly with the oligomer on the support.

Again not wishing to be bound by theory, it is believed that the product ketone from irradiation of a MeNVOC-protected oligomer reacts at a slower rate with nucleophiles on the oligomer than the product aldehydes from irradiation of the same NVOC-protected oligomer. Although not unambiguously determined, it is believed that this difference in reaction rate is due to the difference in general reactivity between aldehydes and ketones towards nucleophiles due to steric and electronic effects.

The photoremovable protecting groups of the present invention are readily removed. For example, the photolysis of N-protected L-phenylalanine in solution having different photoremovable protecting groups was analyzed, and the results are presented in the following table:

Table

Photolysis of Protected L-Phe-OH

	$t_{1/2}$ in seconds			
Solvent	NBOC	NVOC	MeNVOC	MeNPOC
Dioxane	1288	110	24	19
5mM H_2SO_4/Dioxane	1575	98	33	22

The half life, $t_{1/2}$, is the time in seconds required to remove 50% of the starting amount of protecting group. NBOC is the 6-nitrobenzyloxycarbonyl group, NVOC is the 6-nitroveratryloxycarbonyl group, MeNVOC is the methyl-6-nitroveratryloxycarbonyl group, and MeNPOC is the methyl-6-nitropiperonyloxycarbonyl group. The photolysis was carried out in the indicated solvent with 362/364 nm-wavelength irradiation having an intensity of 10 mW/cm^2, and the concentration of each protected phenylalanine was 0.10 mM.

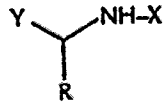
The table shows that deprotection of NVOC-, MeNVOC-, and MeNPOC-protected phenylalanine proceeded faster than the deprotection of NBOC. Furthermore, it shows that the deprotection of the two derivatives that are substituted on the benzylic carbon, MeNVOC and MeNPOC, were photolyzed at the highest rates in both dioxane and acidified dioxane.

1. Use of Photoremovable Groups During Solid-Phase Synthesis of Peptides

The formation of peptides on a solid-phase support requires the stepwise attachment of an amino acid to a substrate-bound growing chain. In order to prevent unwanted polymerization of the monomeric amino acid under the reaction conditions, protection of the amino terminus of the amino acid is required. After the monomer is coupled to the end of the peptide, the N-terminal protecting group is removed, and another amino acid is coupled to the chain. This cycle of coupling and deprotecting is continued for each amino acid in the peptide sequence. See Merrifield, J. Am. Chem. Soc. (1963) 85:2149, and Atherton et al., "Solid Phase Peptide Synthesis" 1989, IRL Press, London, both incorporated herein by reference for all purposes. As described above, the use of a photoremovable protecting group allows removal of selected portions of the

substrate surface, via patterned irradiation, during the deprotection cycle of the solid phase synthesis. This selectively allows spatial control of the synthesis--the next amino acid is coupled only to the irradiated areas.

In one embodiment, the photoremovable protecting groups of the present invention are attached to an activated ester of an amino acid at the amino terminus:



where R is the side chain of a natural or unnatural amino acid, X is a photoremovable protecting group, and Y is an activated carboxylic acid derivative. The photoremovable protecting group, X, is preferably NVOC, NPOC, PyROC, MeNVOC, MeNPOC, and the like as discussed above. The activated ester, Y, is preferably a reactive derivative having a high coupling efficiency, such as an acyl halide, mixed anhydride, N-hydroxysuccinimide ester, perfluorophenyl ester, or urethane protected acid, and the like. Other activated esters and reaction conditions are well known (See Atherton et al.).

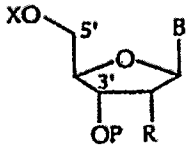
2. Use of Photoremovable Groups During Solid-Phase Synthesis of Oligonucleotides

The formation of oligonucleotides on a solid-phase support requires the stepwise attachment of a nucleotide to a substrate-bound growing oligomer. In order to prevent unwanted polymerization of the monomeric nucleotide under the reaction conditions, protection of the 5'-hydroxyl group of the nucleotide is required. After the monomer is coupled to the end of the oligomer, the 5'-hydroxyl protecting group is removed, and another nucleotide is coupled to the chain. This cycle of

coupling and deprotecting is continued for each nucleotide in the oligomer sequence. See Gait, "Oligonucleotide Synthesis: A Practical Approach" 1984, IRL Press, London, incorporated herein by reference for all purposes. As described above, the use of a photoremovable protecting group allows removal, via patterned irradiation, of selected portions of the substrate surface during the deprotection cycle of the solid phase synthesis. This selectively allows spatial control of the synthesis--the next nucleotide is coupled only to the irradiated areas.

Oligonucleotide synthesis generally involves coupling an activated phosphorous derivative on the 3'-hydroxyl group of a nucleotide with the 5'-hydroxyl group of an oligomer bound to a solid support. Two major chemical methods exist to perform this coupling: the phosphate-triester and phosphoramidite methods (See Gait). Protecting groups of the present invention are suitable for use in either method.

In a preferred embodiment, a photoremovable protecting group is attached to an activated nucleotide on the 5'-hydroxyl group:

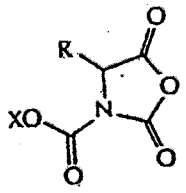


where B is the base attached to the sugar ring; R is a hydrogen atom when the sugar is deoxyribose or R is a hydroxyl group when the sugar is ribose; P represents an activated phosphorous group; and X is a photoremovable protecting group. The photoremovable protecting group, X, is preferably NV, NP, PyR, MeNV, MeNP, NVOC, NPOC, PyROC, MeNVOC, MeNPOC, and the like as described above. The activated phosphorous group, P, is preferably a

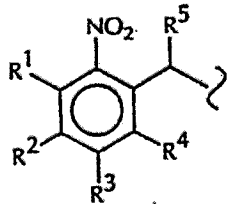
reactive derivative having a high coupling efficiency, such as a phosphate-triester, phosphoramidite or the like. Other activated phosphorous derivatives, as well as reaction conditions, are well known (See Gait).

E. Amino Acid N-Carboxy Anhydrides Protected With a Photoremovable Group

During Merrifield peptide synthesis, an activated ester of one amino acid is coupled with the free amino terminus of a substrate-bound oligomer. Activated esters of amino acids suitable for the solid phase synthesis include halo-formate, mixed anhydride, imidazoyl formate, acyl halide, and also includes formation of the activated ester in situ and the use of common reagents such as DCC and the like (See Atherton et al.). A preferred protected and activated amino acid has the general formula:



where R is the side chain of the amino acid and X is a photoremovable protecting group. This compound is a urethane-protected amino acid having a photoremovable protecting group attached to the amine. A more preferred activated amino acid is formed when the photoremovable protecting group has the general formula:

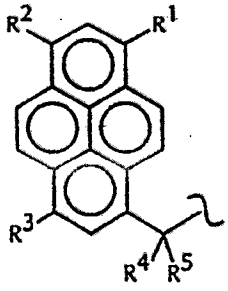


where R^1, R^2, R^3, and R^4 independently are a hydrogen atom, a lower alkyl, aryl, benzyl, halogen, hydroxyl, alkoxyl, thiol, thioether, amino, nitro, carboxyl, formate, formamido or phosphido group, or adjacent substituents (i.e., R^1-R^2, R^2-R^3, R^3-R^4) are substituted oxygen groups

that together form a cyclic acetal or ketal; and R^5 is a hydrogen atom, alkoxyl, alkyl, halo, aryl, or alkenyl group.

A preferred activated amino acid is formed when the photoremovable protecting group is 6-nitroveratryloxycarbonyl. That is, R^1 and R^4 are each a hydrogen atom, R^2 and R^3 are each a methoxy group, and R^5 is a hydrogen atom. Another preferred activated amino acid is formed when the photoremovable group is 6-nitropiperonyl: R^1 and R^4 are each a hydrogen atom, R^2 and R^3 together form a methylene acetal, and R^5 is a hydrogen atom. Other protecting groups are possible. Another preferred activated ester is formed when the photoremovable group is methyl-6-nitroveratryl or methyl-6-nitropiperonyl.

Another preferred activated amino acid is formed when the photoremovable protecting group has the general formula:



where R^1, R^2, and R^3 independently are a hydrogen atom, a lower alkyl, aryl, benzyl, halogen, hydroxyl, alkoxyl, thiol, thioether, amino, nitro, carboxyl, formate, formamido, sulfanate, sulfido or phosphido group, and R^4 and R^5 independently are a hydrogen atom, an alkoxy, alkyl, halo, aryl, or alkenyl group. The resulting compound is a urethane-protected amino acid having a pyrenylmethyloxycarbonyl protecting group attached to the amine. A more preferred embodiment is formed when R^1 through R^5 are each a hydrogen atom.

The urethane-protected amino acids having a photoremovable protecting group of the present invention

are prepared by condensation of an N-protected amino acid with an acylating agent such as an acyl halide, anhydride, chloroformate and the like (See Fuller et al., U.S. Patent No. 4,946,942 and Fuller et al., J. Amer. Chem. Soc. (1990) 112:7414-7416, both herein incorporated by reference for all purposes).

Urethane-protected amino acids having photoremovable protecting groups are generally useful as reagents during solid-phase peptide synthesis, and because of the spatial selectivity possible with the photoremovable protecting groups, are especially useful for the spatially addressing peptide synthesis. These amino acids are difunctional: the urethane group first serves to activate the carboxy terminus for reaction with the amine bound to the surface, and, once the peptide bond is formed, the photoremovable protecting group protects the newly formed amino terminus from further reaction. These amino acids are also highly reactive to nucleophiles, such as deprotected amines on the surface of the solid support, and due to this high reactivity, the solid-phase peptide coupling times are significantly reduced, and yields are typically higher.

IV. Data Collection

A. Data Collection System

Substrates prepared in accordance with the above description are used in one embodiment to determine which of the plurality of sequences thereon bind to a receptor of interest. Fig. 10 illustrates one embodiment of a device used to detect regions of a substrate which contain florescent markers. This device would be used, for example, to detect the presence or absence of a fluorescently labeled receptor such as an antibody which has bound to a synthesized polymer on a substrate.

Light is directed at the substrate from a light source 1002 such as a laser light source of the type well known to those of skill in the art such as a model no.

2025 made by Spectra Physics. Light from the source is directed at a lens 1004 which is preferably a cylindrical lens of the type well known to those of skill in the art. The resulting output from the lens 1004 is a linear beam rather than a spot of light. Thus, data can be detected substantially simultaneously along a linear array of pixels rather than on a pixel-by-pixel basis. It will be understood that while a cylindrical lens is used herein as an illustration of one technique for generating a linear beam of light on a surface, other techniques could also be utilized.

The beam from the cylindrical lens is passed through a dichroic mirror or prism and directed at the surface of the suitably prepared substrate 1008. Substrate 1008 is placed on an x-y translation stage 1009 such as a model no. PM500-8 made by Newport. Certain locations on the substrate will fluoresce and fluorescence will be transmitted along the path indicated by dashed lines back through the dichroic mirror, and focused with a suitable lens 1010 such as an f/1.4 camera lens on a linear detector 1012 via a variable f stop focusing lens 1014. Through use of a linear light beam, it becomes possible to generate data over a line of pixels (such as about 1 cm) along the substrate, rather than from individual points on the substrate. In alternative embodiments, light is directed at a 2-dimensional area of the substrate and fluorescence is detected by a 2-dimensional CCD array. Linear detection is preferred because substantially higher power densities are obtained.

Detector 1012 detects the amount of fluorescence emitted from the substrate as a function of position. According to one embodiment the detector is a linear CCD array of the type commonly known to those of skill in the art. The x-y translation stage, the light source, and the detector 1012 are all operably connected to a computer 1016 such as an IBM PC-AT or equivalent for

control of the device and data collection from the CCD array.

In operation, the substrate is appropriately positioned by the translation stage. The light source is then illuminated, and fluorescence intensity data are gathered with the computer via the detector.

In an alternate embodiment, the substrate and x/y translation table are placed under a microscope which includes one or more objectives. Light (about 488 nm) from a laser, which in some embodiments is a model no. 2020-05 argon ion laser manufactured by Spectraphysics, is directed at the substrate by a dichroic mirror which passes greater than about 520 nm light but reflects 488 nm light. The dichroic mirror may be, for example, a model no. FT510 manufactured by Carl Zeiss. Light reflected from the mirror then enters the microscope which may be, for example, a model no. Axioscop 20 manufactured by Carl Zeiss. Fluorescein-marked materials on the substrate will fluoresce >488 nm light, and the fluoresced light will be collected by the microscope and passed through the mirror. The fluorescent light from the substrate is then directed through a wavelength filter and, thereafter through an aperture plate. The wavelength filter may be, for example, a model no. OG530 manufactured by Melles Griot and the aperture plate may be, for example, a model no. 477352/477380 manufactured by Carl Zeiss.

The fluoresced light then enters a photomultiplier tube which in some embodiments is a model no. R943-02 manufactured by Hamamatsu, the signal is amplified in a preamplifier and photons are counted by a photon counter. The number of photons is recorded as a function of the location in the computer. The pre-amp may be, for example, a model no. SR440 manufactured by Stanford Research Systems and the photon counter may be a model no. SR400 manufactured by Stanford Research Systems. The substrate is then moved to a subsequent